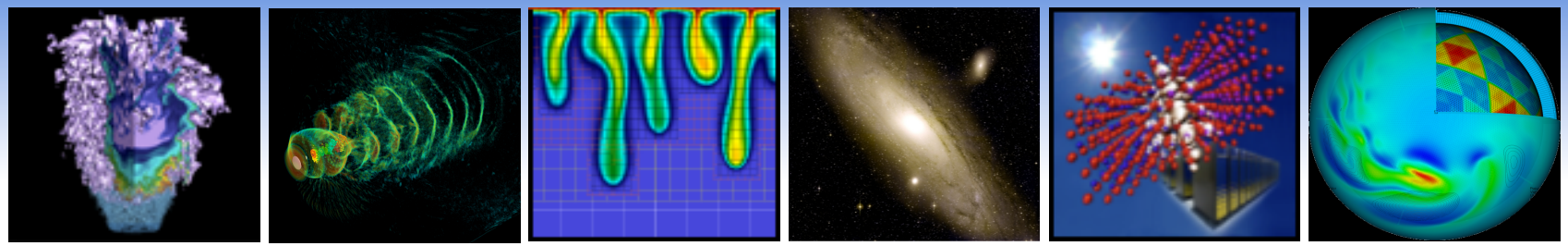


Debunking some Common Misconceptions of Science in the Cloud

Shane Canon
Lawrence Berkeley National Lab

ScienceCloud 2011
San Jose, CA

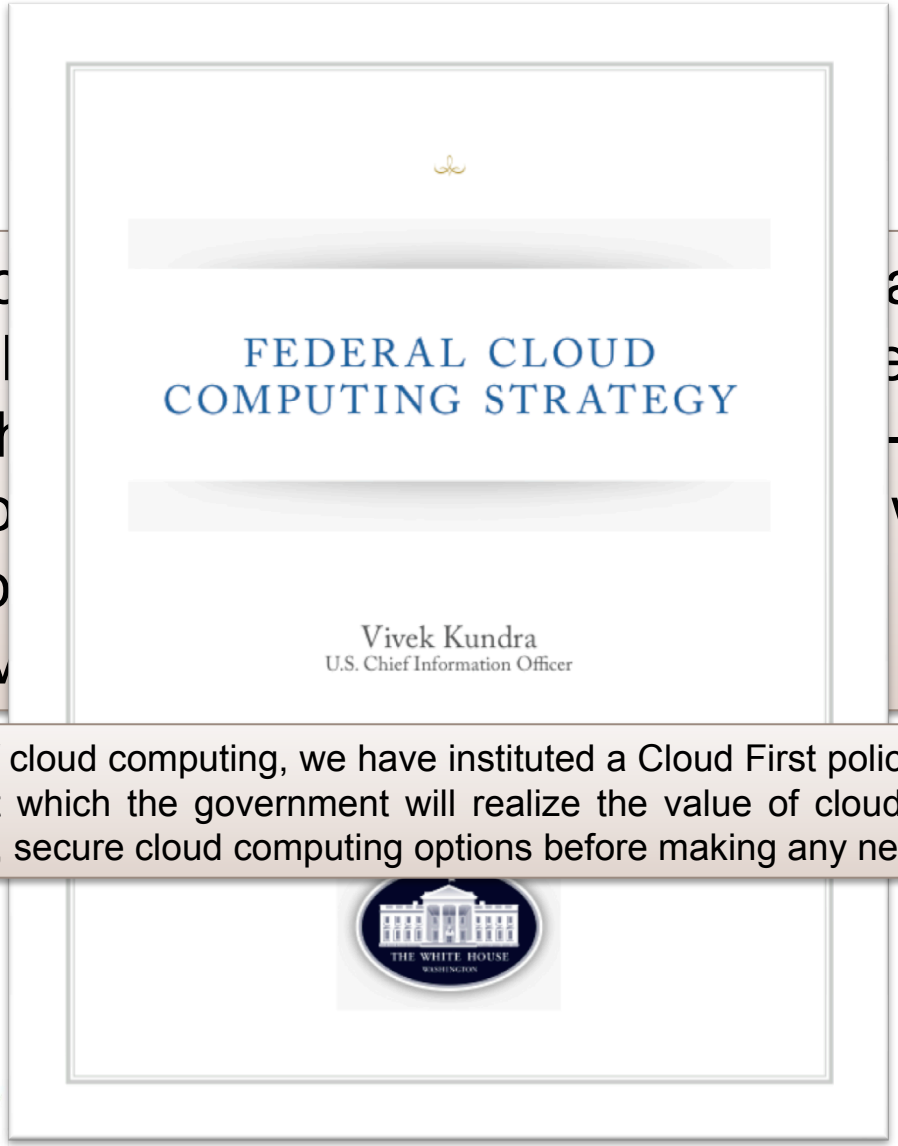
U.S. DEPARTMENT OF ENERGY | Office of Science

Berkeley Lab — 75 Years of World-Class Science 1931-2006

The Push towards Clouds

A survey c
the world
average) h
hosted clo
year (28 p
--Ovum (M

anies across
ercent on
-party-
with last



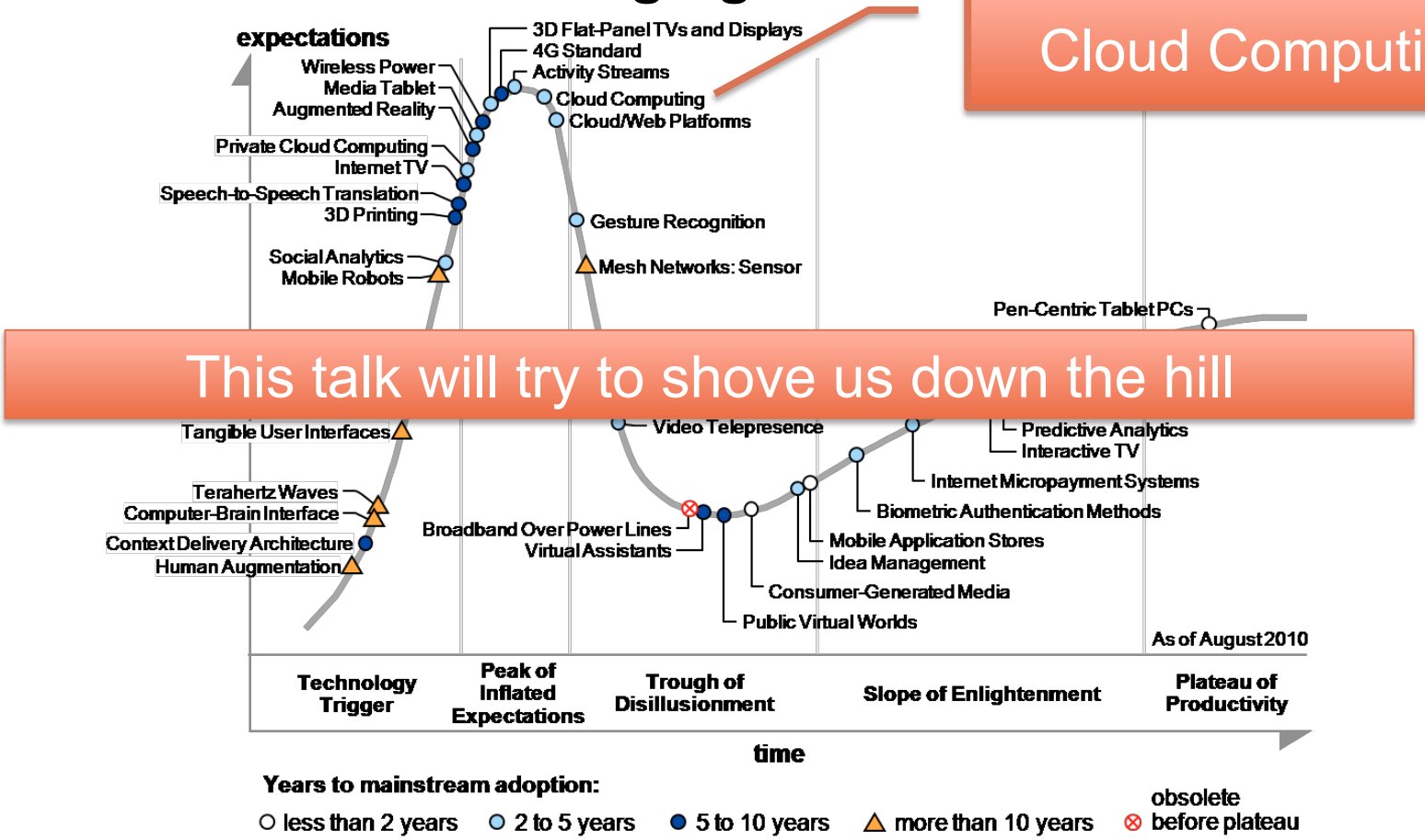
To harness the benefits of cloud computing, we have instituted a Cloud First policy. This policy is intended to accelerate the pace at which the government will realize the value of cloud computing by requiring agencies to evaluate safe, secure cloud computing options before making any new investments

The Hype around Clouds

Gartner's 2010 Emerging Technologies Hype Cycle

Cloud Computing

This talk will try to shove us down the hill



Cloud Misconceptions

- Clouds are simple to use and don't require system administrators.
- My job will run immediately in the cloud.
- Clouds are more efficient.
- Clouds allow you to *ride* Moore's Law without additional investment.
- Commercial Clouds are much cheaper than operating your own system.

Are Clouds Easy to Use?

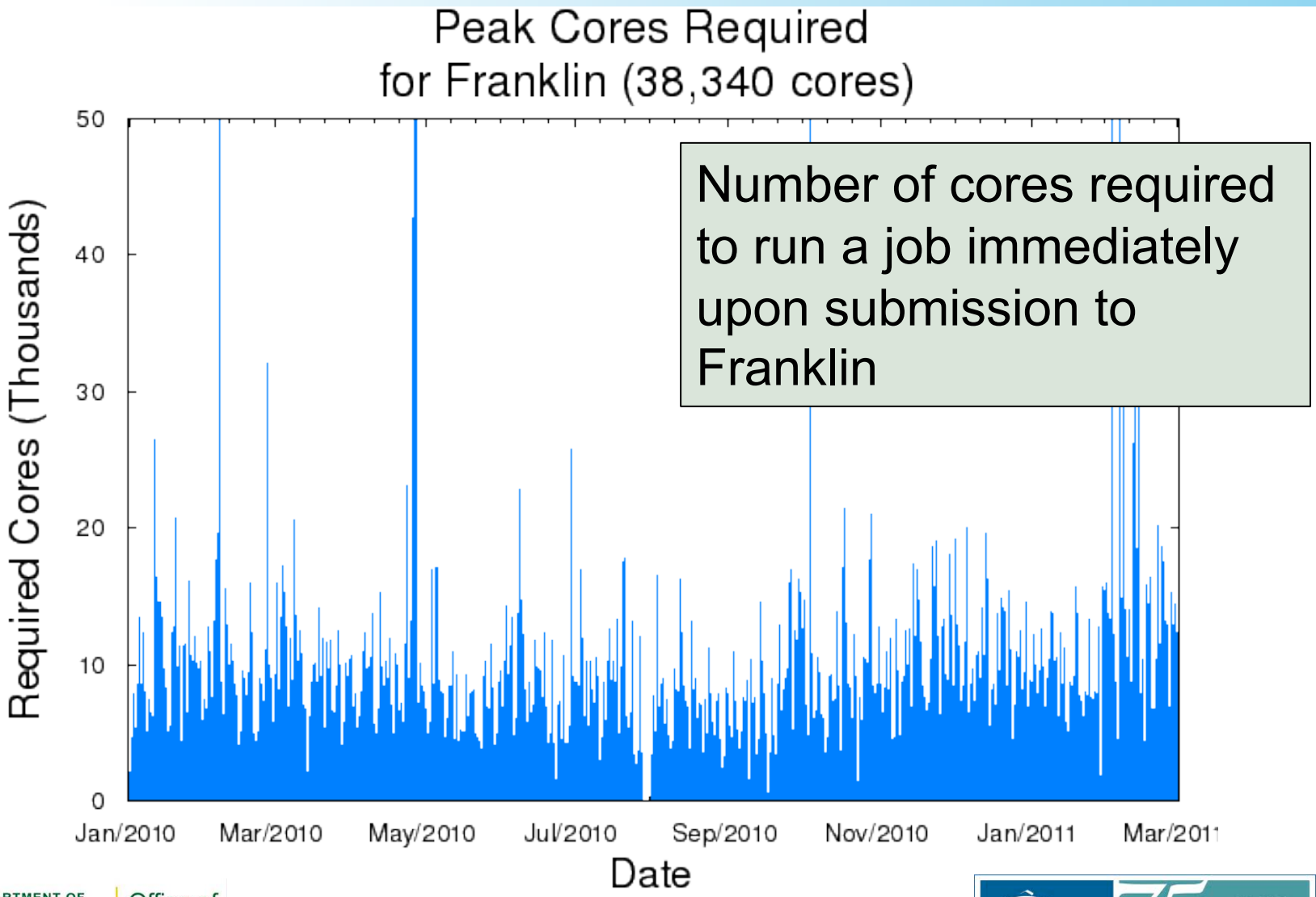
From Experience with Magellan we have Learned

- **IaaS Clouds can require significant amounts of system administration expertise**
- **Images must be customized for the application**
- **No batch environment. No global file system.**
- **Users must properly secure and protect their images and instances.**
- **Do we want to turn scientists into system administrators?**

Cloud Misconceptions

- Clouds are simple to use and don't require system administrators.
- My job will run immediately in the cloud.
- Clouds are more efficient.
- Clouds allow you to *ride* Moore's Law without additional investment.
- Commercial Clouds are much cheaper than operating your own system.

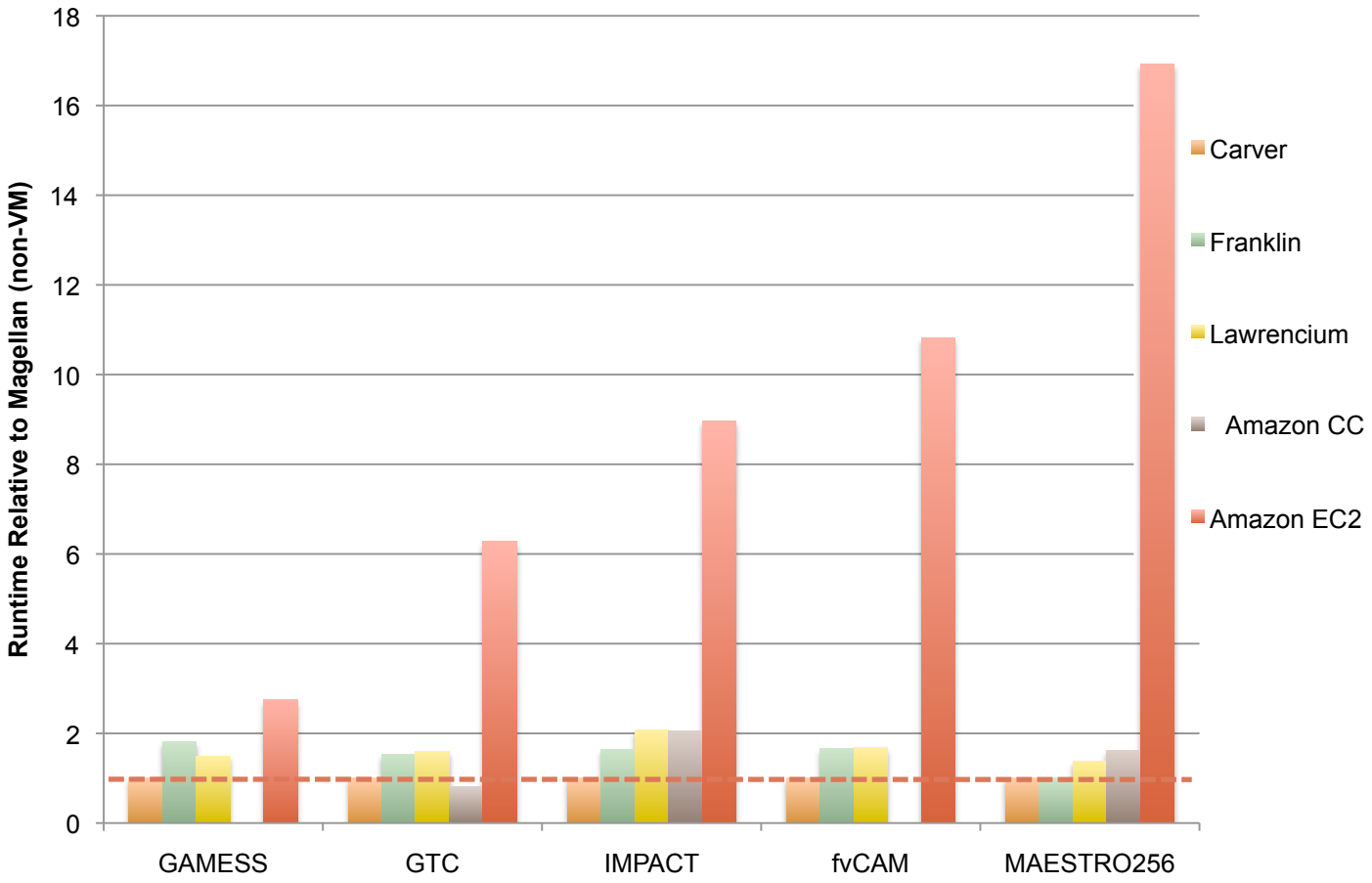
Is the Cloud Elastic enough for HPC?



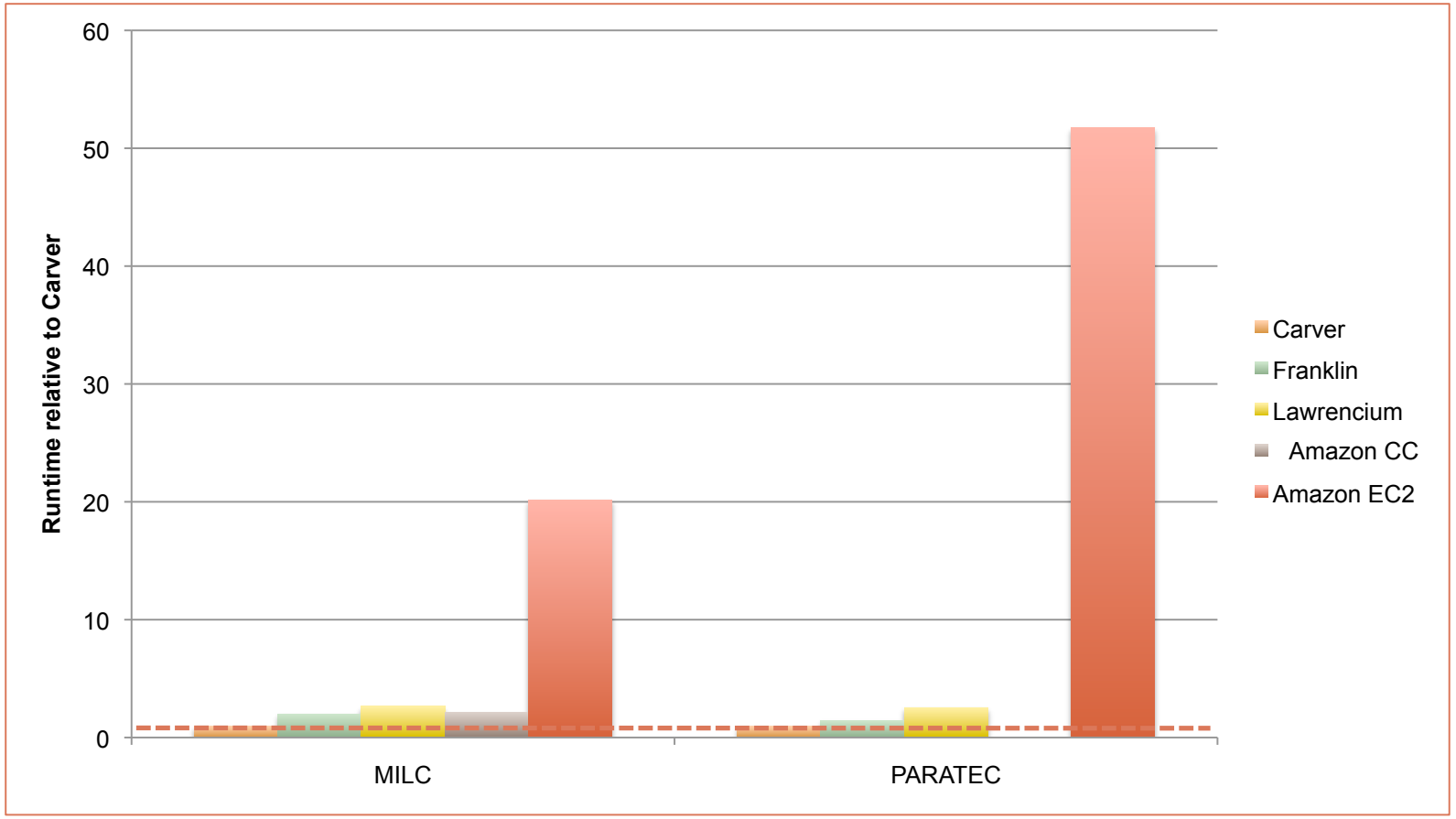
Cloud Misconceptions

- Clouds are simple to use and don't require system administrators.
- My job will run immediately in the cloud.
- Clouds are more efficient.
- Clouds allow you to *ride* Moore's Law without additional investment.
- Commercial Clouds are much cheaper than operating your own system.

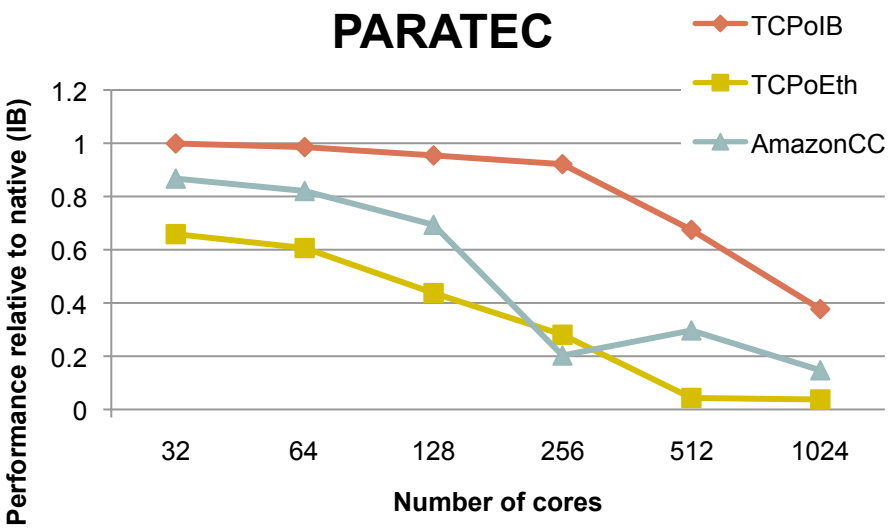
Application Performance
Application Benchmarks

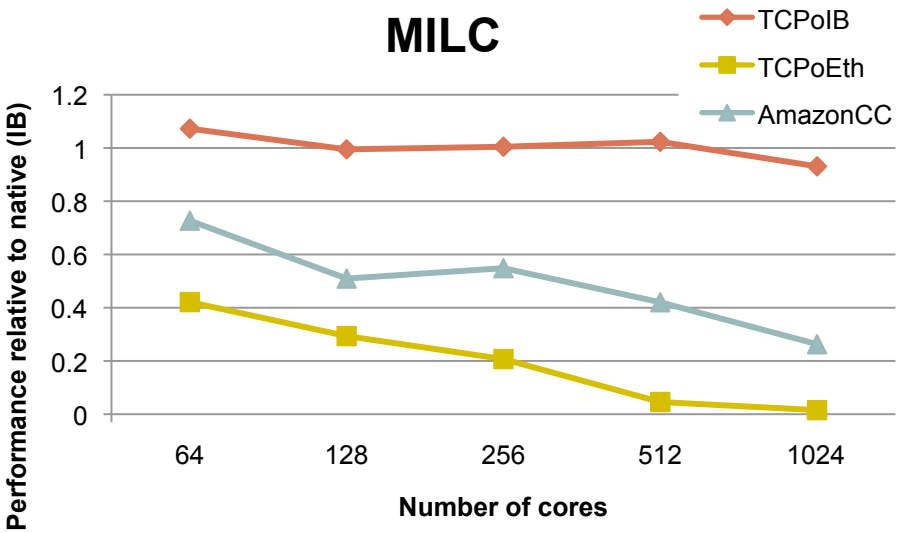


Application Performance
Application Benchmarks



Application Scaling





Assumptions for cost saving from Clouds aren't true for HPC Centers.

EFFICIENCY	
Cloud Benefits	**Current Environment**
• Improved asset utilization (server utilization > 60-70%) • Aggregated demand and accelerated system consolidation (e.g., Federal Data Center Consolidation Initiative) • Improved productivity in application development, application management, network, and end-user	• Low asset utilization (server utilization < 30% typical) • Fragmented demand and duplicative systems • Difficult-to-manage systems
AGILITY	
Cloud Benefits	**Current Environment**
• Purchase "as-a-service" from trusted cloud providers • Near-instantaneous increases and reductions in capacity • More responsive to urgent agency needs	• Years required to build data centers for new services • Months required to increase capacity of existing services
INNOVATION	
Cloud Benefits	**Current Environment**
• Shift focus from asset ownership to service management • Tap into private sector innovation • Encourages entrepreneurial culture • Better linked to emerging technologies (e.g., devices)	• Burdened by asset management • De-coupled from private sector innovation engines • Risk-adverse culture

- HPC Centers run at >90% CPU utilization and >90% scheduled utilization.
- HPC Centers partner with Vendors to field cutting edge systems
- HPC more aggressive with technical risks

From the Federal Cloud Computing Strategy

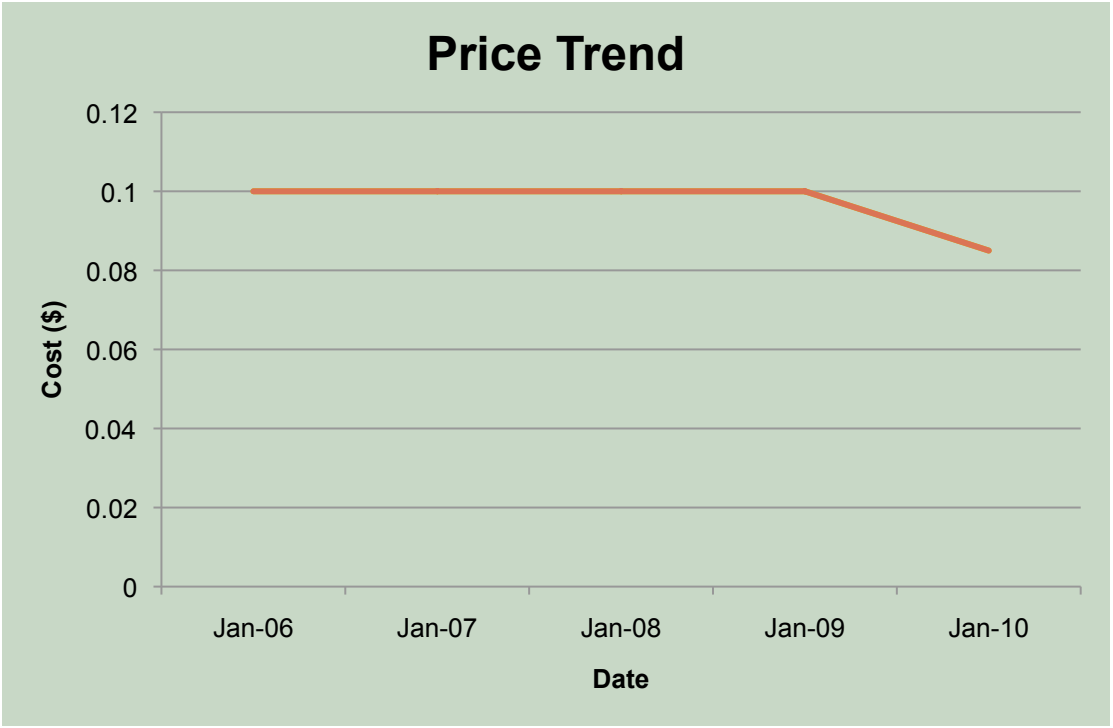
Enterprise IT versus HPC

	Traditional Enterprise IT	HPC Centers
Typical Load Average	30% *	90%
Computational Needs	Bounded computing requirements – Sufficient to meet customer demand or transaction rates. (i.e. If you gave a typical business free computing, would they suddenly be able to take advantage of it?)	Virtually unbounded requirements – Scientist always have larger, more complicated problems to simulate or analyze.
Scaling Approach	Scale-in. Emphasis on consolidating in a node using virtualization	Scale-Out Applications run in parallel across multiple nodes.

Cloud Misconceptions

- Clouds are simple to use and don't require system administrators.
- My job will run immediately in the cloud.
- Clouds are more efficient.
- Clouds allow you to *ride* Moore's Law without additional investment.
- Commercial clouds are much cheaper than operating your own system.

Cloud Pricing Trends



The cost of a standard cloud instance has dropped 18% over 5 years. Meanwhile, cores per socket have increased 2x-5x per socket in the same time-frame at roughly constant cost.

Cloud Misconceptions

- Clouds are simple to use and don't require system administrators.
- My job will run immediately in the cloud.
- Clouds are more efficient.
- Clouds allow you to *ride* Moore's Law without additional investment.
- Commercial Clouds are much cheaper than operating your own system.

Cost of NERSC in the Cloud

Component	Cost
Compute Systems (1.38B hours)	$180,900,000
HPSS (17 PB)	$12,200,000
File Systems (2 PB)	$2,500,000
Total (Annual Cost)	**$195,600,000**

Assumes 85% utilization and zero growth in HPSS and File System data.
Doesn't include the 2x-10x performance impact that has been measured.
This still only captures about 65% of NERSC's $55M annual budget.
No consulting staff, no administration, no support.

Where are (commercial) clouds effective?

- Individual projects with high-burst needs.
  - Avoid paying for idle hardware
  - Access to larger scale (elasticity)
  - Alternative: Pool with other users (condo model)
- High-Throughput Applications with modest data needs
  - Bioinformatics
  - Monte-Carlo simulations
  - Cost issues still apply
- Infrastructure Challenged Sites
  - Facilities cost >> IT costs
  - Consider the long-term costs
- Undetermined or Volatile Needs
  - Use Clouds to baseline requirements and build in-house

Is an HPC Center a Cloud?

HPC Centers ?

- ***Resource pooling.*** ✓
- ***Broad network access.*** ✓
- ***Measured Service.*** ✓
- ***Rapid elasticity.***
  - **Usage can grow/shrink; pay-as-you-go.** ✓
- ***On-demand self-service.***
  - **Users cannot demand (or pay for) more service than their allocation allows** X
  - **Jobs often wait for hours or days in queues**

From the NIST definition for Cloud Computing

It's All Business

- Cloud computing is a business model
- It can be applied to HPC systems as well as traditional clouds
- Achieve on-demand elasticity through:
  - Idle hardware (at ownership cost)
  - Sharing cores/nodes (at performance cost)
  - Scheduling policies (pre-emption)

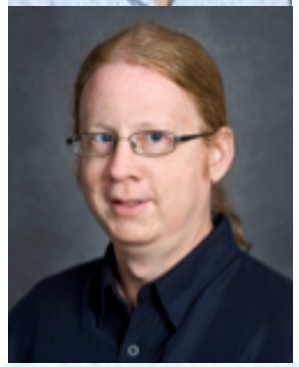
Closing Remarks

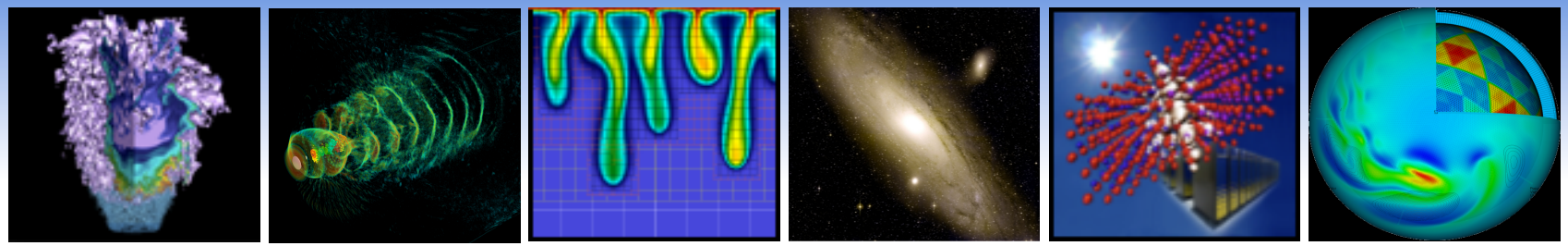
- **Cloud Computing as it exist today is not ready for HPC**
  - **Overhead to convert to Cloud environments**
  - **Virtual instances underperform bare-metal systems**
  - **Less cost-effective than most large centers**
- **Traditional HPC resource providers can learn from the Cloud (Magellan presentation will discuss this)**

Acknowledgements

- Lavanya Ramakrishnan
- Iwona Sakrejda
- Tina Declerck
- Others
  - Keith Jackson
  - Nick Wright
  - John Shalf
  - Krishna Muriki (not picture)





Thank you!

Contact Info:
Shane Canon
Scanon@lbl.gov
magellan.nersc.gov

U.S. DEPARTMENT OF ENERGY | Office of Science

Berkeley Lab — 75 Years of World-Class Science 1931-2006